AXIOM OF CHOICE

Eric Schechter, Vanderbilt University

The Axiom of Choice (**AC**) was formulated about a century ago, and it was controversial for a few decades after that; it may be considered the last great controversy of mathematics. It is now a basic assumption used in many parts of mathematics. In fact, assuming AC is equivalent to assuming any of these principles (and many others):

- Given any two sets, one set has cardinality less than or equal to that of the other set -- i.e., one set is in one-to-one correspondence with some subset of the other. (*Historical remark:* It was questions like this that led to Zermelo's formulation of AC.)
- Any vector space over a field F has a basis -- i.e., a maximal linearly independent subset -- over that field. (*Remark:* If we only consider the case where F is the real line, we obtain a slightly weaker statement; it is not yet known whether this statement is also equivalent to AC.)
- Any product of compact topological spaces is compact. (This is now known as *Tychonoff's Theorem*, though Tychonoff himself only had in mind a much more specialized result that is *not* equivalent to the Axiom of Choice.)

AC has many forms; here is one of the simplest:

> **Axiom of Choice.** Let C be a collection of nonempty sets. Then we can *choose* a member from each set in that collection. In other words, there exists a function f defined on C with the property that, for each set S in the collection, $f(S)$ is a member of S.

The function f is then called a *choice function*.

To understand this axiom better, let's consider a few examples.

- If C is the collection of all nonempty subsets of $\{1,2,3,...\}$, then we can define f quite easily: just let $f(S)$ be the smallest member of S.
- If C is the collection of all intervals of real numbers with positive, finite lengths, then we can define $f(S)$ to be the midpoint of the interval S.
- If C is some more general collection of subsets of the real line, we may be able to define f by using a more complicated rule.
- However, if C is the collection of *all* nonempty subsets of the real line, it is not clear how to find a suitable function f. In fact, no one has ever found a suitable function f for this collection C, and there are convincing model-theoretic arguments that no one ever will. (Of course, to *prove* this requires a precise definition of "find," etc.)

The controversy was over how to interpret the words "choose" and "exists" in the axiom:

- If we follow the constructivists, and "exist" means "find," then the axiom is *false*, since we cannot find a choice function for the nonempty subsets of the reals.
- However, most mathematicians give "exists" a much weaker meaning, and they consider the Axiom to be true: To define $f(S)$, just arbitrarily "pick any member" of S.

In effect, when we accept the Axiom of Choice, this means we are agreeing to the convention that we shall permit ourselves to use a hypothetical choice function f in proofs, as though it "exists" in some sense, even in cases where we cannot give an explicit example of it or an explicit algorithm for it.

The "existence" of f -- or of any mathematical object, even the number "3" -- is purely formal. It does not have the same kind of solidity as your table and your chair; it merely exists in the mental universe of mathematics. Many different mathematical universes are possible. When we accept or reject the Axiom of Choice, we are specifying something about which mental universe we're choosing to work in. Both possibilities are feasible -- i.e., neither accepting nor rejecting AC yields a contradiction; that fact follows from models devised by Gödel and Cohen. However, most "ordinary" mathematicians -- i.e., most mathematicians who are not logicians or set theorists -- accept the Axiom of Choice chiefly because their work is *simpler* with the Axiom of Choice than without it.

Bertrand Russell was more famous for his work in philosophy and political activism, but he was also an accomplished mathematician. His book *Introduction to Mathematical Philosophy* includes some discussion of AC. Here is my paraphrasing of part of what he said:

To choose one sock from each of infinitely many pairs of socks requires the Axiom of Choice, but for shoes the Axiom is not needed.

The idea is that the two socks in a pair are identical in appearance, and so we must make an arbitrary choice if we wish to choose one of them. For shoes, we can use an explicit algorithm -- e.g., "always choose the left shoe." Why does Russell's statement mention *infinitely many* pairs? Well, if we only have finitely many pairs of socks, then AC is not needed -- we can choose one member of each pair using the definition of "nonempty," and we can repeat an operation finitely many times using the rules of formal logic (not discussed here).

A few pure mathematicians and many applied mathematicians (including, e.g., some mathematical physicists) are uncomfortable with the Axiom of Choice. Although AC simplifies some parts of mathematics, it also yields some results that are unrelated to, or perhaps even contrary to, everyday "ordinary" experience; it implies the existence of some rather bizarre, counterintuitive objects. Perhaps the most bizarre is the **Banach-Tarski Paradox**: It is possible to take the 3-dimensional closed unit ball,

$$B \quad = \quad \left\{ (x,y,z) \in \mathbf{R}^3 : x^2 + y^2 + z^2 \leq 1 \right\}$$

and partition it into finitely many pieces, and move those pieces in rigid motions (i.e., rotations and translations, with pieces permitted to move through one another) and reassemble them to form *two* copies of *B*.

At first glance, the Banach-Tarski result seems to contradict some of our intuition about physics -- e.g., the Law of Conservation of Mass, from classical Newtonian physics. If we assume that the ball has a uniform density, then the Banach-Tarski Paradox seems to say that we can disassemble a one-kilogram ball into pieces and rearrange them to get two one-kilogram balls. But actually, the contradiction can be explained away: Only a set with a defined *volume* can have a defined mass. A "volume" can be defined for many subsets of $\mathbf{R}^3$ --- spheres, cubes, cones, icosahedrons, etc. --- and in fact a "volume" can be defined for nearly any subset of $\mathbf{R}^3$ that we can think of. This leads beginners to expect that the notion of "volume" is applicable to *every* subset of $\mathbf{R}^3$. But it's not. In particular, the pieces in the Banach-Tarski decomposition are sets whose volumes cannot be defined.

Personally, I am not surprised to find the Axiom of Choice coming into play in a subject that is so inherently complicated as unmeasurable sets. I am much more surprised to find AC coming into play in this simpler and more concrete example: I want to classify all subsets of {1, 2, 3, 4, 5, . . .} as "small" or not "small," defining the word "small" in such a way that

a. any set with zero or one members is "small";
b. any union of two "small" sets is "small"; and
c. a set is"small" if and only if its complement isn�t "small."

Now, without much difficulty I can give examples satisfying any *two* of those three rules:

- Define "small" to mean "finite." This satisfies rules a and b. But it does not satisfy rule c, since the even numbers and the odd numbers are complements of each other, and neither of those sets is finite.
- Say that a set is "small" if the number 1 is not a member of that set. This definition satisfies rules b and c, but it classifies the set {1} as "not small," thus failing rule a.
- Say that a set is "small" if it contains at most one of the three numbers 1, 2, 3. That satisfies rules a and c. But it classifies the sets {1} and {2} as small and the set {1,2} as not small, thus failing rule b.

Does there exist a classification scheme satisfying *all three* rules? It turns out that such a classification scheme exists, but an *example* of such a classification scheme does *not* exist (which makes it a bit hard to visualize!). And by that I do not mean just that we haven't found an example yet. I mean that the proofs of existence are *inherently nonconstructive* -- i.e., they cannot be replaced by constructive proofs -- so no examples can *ever* be given. But the proof of that fact is very deep, and it raises interesting philosophical questions: In what sense does that classification scheme "exist"? (My own attitude is that I'm not really working with the classification schemes themselves; I'm just working with *sentences* about *hypothetical* classification schemes.)